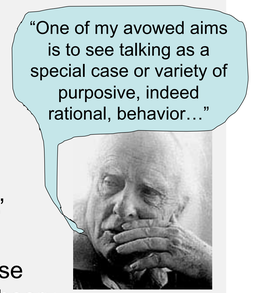
Predicting object and scene descriptions with an information-theoretic model of pragmatics

Michael C. Frank, Avril Kenney, Noah D. Goodman, Joshua Tenenbaum, Antonio Torralba, & Aude Oliva

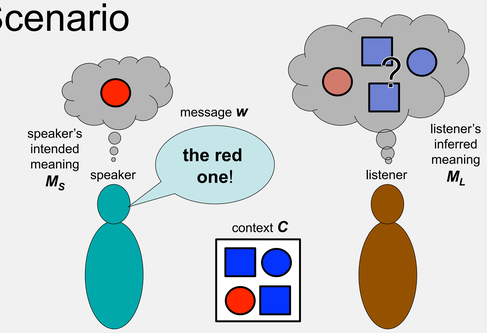
Formalizing Grice's maxims

- Normative maxims:
  - *Quantity*: Be informative
  - *Quality*: Be truthful
  - *Relation*: Be relevant
  - *Manner*: Be perspicuous
- Used by listeners to make inferences about speakers' intentions
- Our goal is to formalize these maxims using information theory



Grice (1975)

Scenario

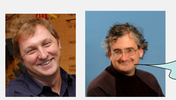


KL Divergence as measure

distribution	probabilities	$D_{KL}(M_S \| w)$
speaker's intended meaning M_S	1	0.00 bits
possible meanings for listener M_L	.25 .25 .25 .25	2.00 bits
extension of "circle"	.5 .5	1.00 bits
extension of "red"	1	0.00 bits

$$p(w|M_S, C) \propto e^{-D_{\mathrm{KL}}(M_S \| w)}$$

$$\propto \frac{1}{|w|}$$

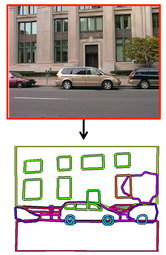


Dale & Reiter (1996)

"Grice's maxims taken collectively mean 'Don't include elements that don't do anything.' Under a goal-oriented view of language generation, there is no need to explicitly follow such a directive at all; the desired behaviour just falls out of the mechanism."

Central claim: By assuming that speakers attempt to communicate optimally in context, listeners and learners can infer meanings even from ambiguous messages.

Scene descriptions



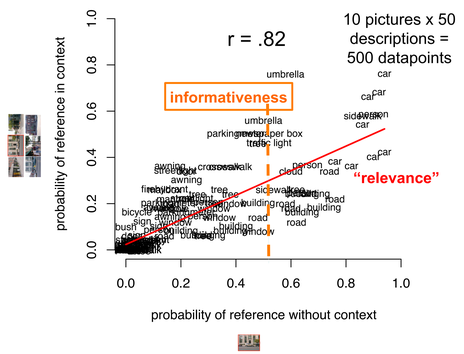
hand-labeled object boundaries (Russell, Torralba, et al., 2008)

- Task: choose objects to pick a scene out of a set of contexts
- Goal: predict which words are chosen using informativeness model
- LabelMe (online database of hand-segmented images) provides ground truth
- Analysis: match descriptions to objects, calculate probability of referring to a particular object.

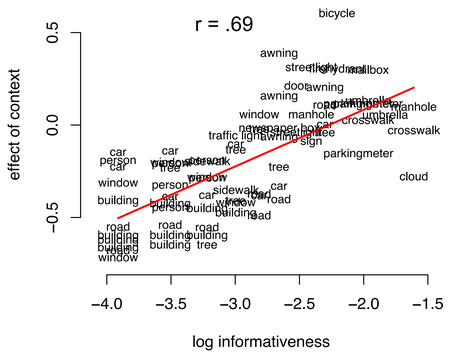
Context condition

No Context condition

Context: relevance and informativeness

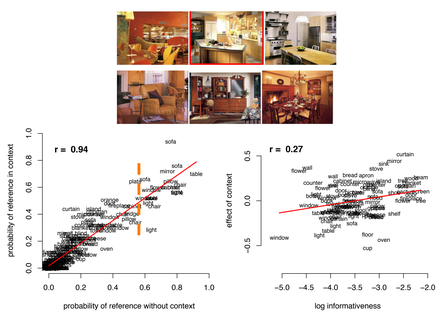


Residual context effects vs. informativeness



Subordinate use in context

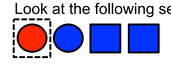
Indoor scenes had a smaller context effect



Object descriptions

Word learning

Look at the following set of objects:



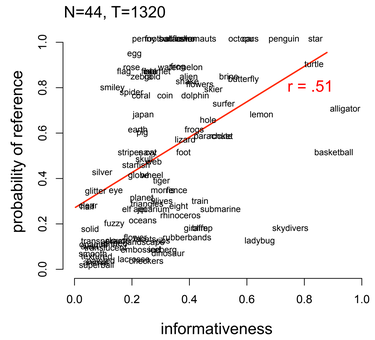
How many red objects are there?
How many circular objects are there?

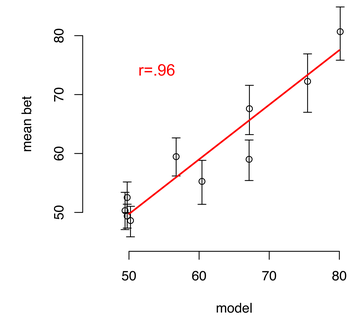
Now imagine someone is talking to you in a foreign language. You don't know the meaning of the adjective, *daxy*, that he uses to refer to the object with the box around it.

Your job is to guess the meaning of *daxy*. Your guess should take the form of "bets." Imagine that you have $100 to spend betting on the meaning of the word. You should divide your money among the possible meanings -- the amount of money bet on each option should correspond to how confident you are that it is correct. Bets must sum to 100!

The one with the box around it is *daxy*. What do you think *daxy* means?

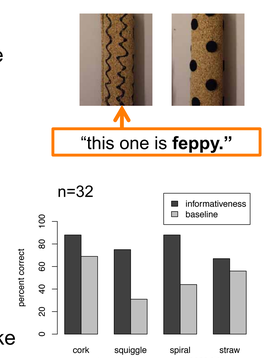
red: _____ circular: _____





Informativeness inferences in children

- 3-4 year olds
- Novel substance and texture properties
  - counterbalance which one is named
- Two trials/child
- Conditions:
  - context
  - baseline
- Test whether children can make these inferences



Conclusions

- Tested predictions of communication framework
  - more realistic stimuli: real-world objects and scenes
  - more natural response format: keywords and sentences
- Hypothesize features are chosen by the product of two things
  - How relevant they are
  - How informative they are in context
- Factoring this product is often difficult
  - When we can measure each factor independently we can predict responses